

DOCUMENT RESUME

ED 352 026

IR 015 864

AUTHOR MacFarland, Thomas W.
 TITLE Statistics: Notes and Examples. Study Guide for the
 Doctor of Arts in Computer-Based Learning.
 INSTITUTION Nova Univ., Fort Lauderdale, FL. Center for Computer
 and Information Sciences.
 PUB DATE [89]
 NOTE 100p.; Document text is entirely in handwritten
 script.
 PUB TYPE Guides - Classroom Use - Instructional Materials (For
 Learner) (051)
 EDRS PRICE MF01/PC04 Plus Postage.
 DESCRIPTORS Analysis of Variance; Chi Square; *Computer Oriented
 Programs; *Computer Science Education; Correlation;
 Graduate Study; Higher Education; *Statistical
 Analysis; Statistical Significance; *Statistics;
 Study Guides

ABSTRACT

This study guide presents lessons on hand calculating various statistics: Central Tendency and Dispersion; Tips on Data Presentation; Two-Tailed and One-Tailed Tests of Significance; Error Types; Standard Scores; Non-Parametric Tests such as Chi-square, Spearman Rho, Sign Test, Wilcoxon Matched Pairs, Mann-Whitney U, Kruskal-Wallis, and Rank Sums; and Parametric Tests such as correlation and regression, student's t-test, analysis of variance: simple ratio, one-way, and two-way, and tips on using ANOVA and the F-test. (ALF)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED352026

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☐ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Doctor of Arts Computer-Based Learning

STATISTICS: NOTES AND EXAMPLES

Study Guide
Educational Text Series
Nova University Press



"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
Thomas MacFarland

NOVA
UNIVERSITY

Center for Computer - Based Learning
3301 College Avenue
Fort Lauderdale, Florida 33314

BEST COPY AVAILABLE

12015864

STATISTICS : NOTES & EXAMPLES

Thomas W. MacFarland
(tommac)

CENTRAL TENDENCY
AND

DISPERSION

MEASURES OF Central Tendency

~ a measure of the central location of frequency distributions

Mode: the most frequently occurring datum in a set of data

there may be one or more modes

Median: the mid-point in an array (set of data)

half of all "scores" fall below the median and half of all "scores" are above the median

Mean: the arithmetic average

$$\frac{\sum X_i}{N} = \bar{X} \quad (X \text{ bar})$$

The mean is often (but - not always) the preferred statistic when analyzing central tendency

1. The MEAN is the element / point where the sum of all deviations = 0.

2. The mean is affected by extreme scores.

3. MEAN AND standard deviation ($\bar{X}$ & SD) is typical to parametric analysis.

4. Median is typical to non-parametric analysis.

central tendency - 01

Example 1: Mode

2 4 1 2 4 2 3 6 1 7 6
3 2 5 2 1 2 8 1 3 5 2

$$f_1 = 4$$

$$f_5 = 2$$

$$f_2 = 7$$

$$f_6 = 2$$

$$\text{Mode} = 2$$

$$f_3 = 3$$

$$f_7 = 1$$

$$f_4 = 2$$

$$f_8 = 1$$

Example 2: Median

8
7
6
6
5
5
4
4
3
3
3
3
2
2
2
2
2
2
2
2
1
1
1
1

$$\text{Median Even} = \frac{\frac{1}{2}x + \frac{1}{2}y}{2} = \frac{3+2}{2}$$

$$\text{Median Odd} = \text{middle point}$$

$$\frac{3+2}{2} = 2.5$$

$$\text{Median} = 2.5$$

central tendency 01-02

Example 3: Mean

8
7
6
6
6
5
5
5
4
4
3
3
3
3
2
2
2
2
2
2
2
1
1
1
1
1

$$\bar{X} = \frac{\sum X_i}{N} = \frac{72}{22} = 3.27$$

$$V = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N-1} = \frac{330 - \frac{(72)^2}{22}}{22-1}$$

$$V = \frac{330 - \frac{5184}{22}}{21} = \frac{330 - 235.64}{21}$$

$$V = \frac{94.36}{21} = 4.49$$

$$SD = \sqrt{V}$$

$$SD = \sqrt{4.49}$$

$$SD = 2.12 \quad \rightsquigarrow \text{N-1 Method}$$

$$\bar{X} = 3.27$$

$$SD = 2.12$$

Dispersion 01-02

Variance in Opinion : N or N-1

N ... large samples (population)

N-1 ... small samples (N < 30)

$$SD = 2.07 \quad (N)$$

$$SD = 2.12 \quad (N-1)$$

Variance: the sum of the squared deviations from the mean. DIVIDED by N

Standard Deviation: square root of the variance

Even when median is reported, it is still typical to report mean and SD.

SD:

1. A reflection of the Dispersion of scores
2. Allows for exact interpretation of scores
3. A central point in many inferential analyses

Use of $N-1$: preferred for small samples

Use of N : " " population

Generally:

$N-1$ if $n < 30$

N if $n \geq 30$

Note: $N-1$ or N basically a moot point for analysis with a large N

Tips on Data Presentation

Reporting N, Percent of Total, $\bar{X}$, and SD

- a common technique in the presentation of data is the practice of reporting data in tabular form

DIS	4 5 2 1 4 5 3 2 1	n = 9	n/N = 22.50%
DIS _c	5 2 1 3 3 2 4 2 1	n = 9	n/N = 22.50%
DTL	4 3 1 4 3 2 5 5 4	n = 9	n/N = 22.50%
CEJ	4 2 3 4 5 1	n = 7	n/N = 17.50%
MS	2 3 5 2 1 3	n = 6	n/N = 15.00%

TABLE 1

Response to Surveys by CCIS/CBL Program

Program	N	% of Total	$\bar{X}$	SD
CEJ	7	17.50	3.40	1.51
DIS	9	22.50	3.00	1.58
DIS _c	9	22.50	2.56	1.33
DTL	9	22.50	3.44	1.33
MS	6	15.00	2.67	1.37
TOTAL	40	100.00	3.03	1.41

NOTE. SD is based on the N method.

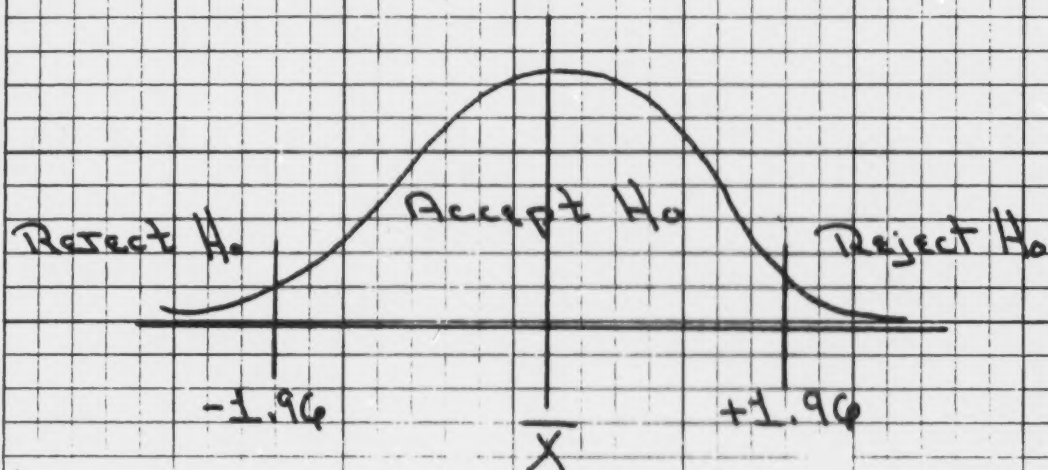
Reporting 01-01

TWO-TAILED AND ONE-TAILED

TESTS OF SIGNIFICANCE

Two-Tailed Test of Significance

H_0 : There is no difference between X and Y in terms of criterion Z .



$$t_{(\alpha=.05, \infty df)} = 1.960 \text{ (two-tailed)}$$

Two-Tailed test is used when H_0 is stating as NO difference.

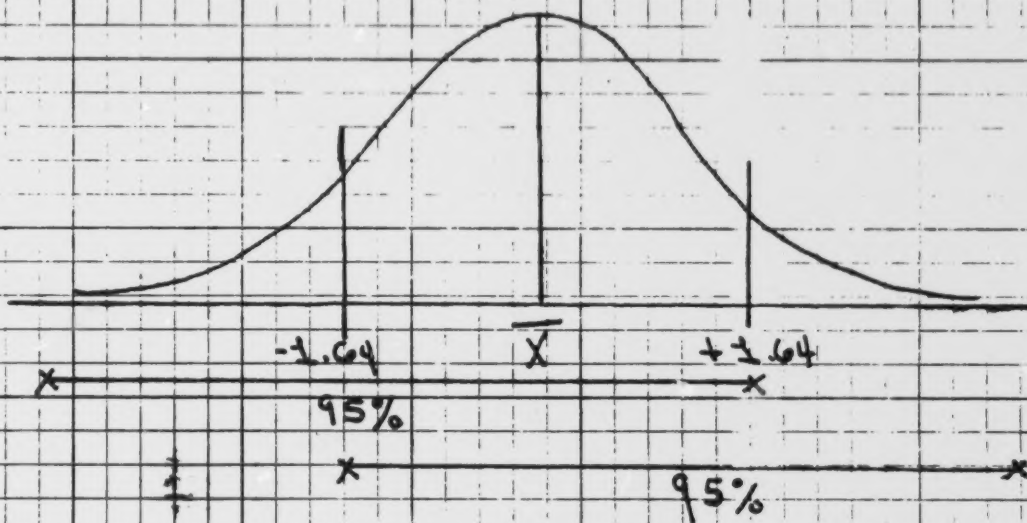
One-Tailed Test of Significance

H_0 : $\bar{X}$ is greater (or less) than μ

H_a : in terms of criterion Z

This a hypothesis of Direction,
not "no Difference."

- In a one-tailed test, the α area of rejection is at either end of the curve
- the t value is less since it is not necessary to go to the extreme end(s) to reach the area of rejection



$$t(\alpha=.05, df=\infty) = 1.645 \text{ (one-tailed)}$$

ERROR TYPES

Error Types

Type I Error

- reject H_0 when it is actually true
- to claim that there is a Difference (accept H_A), when there really is no Difference

Type II Error

- accept H_0 when it is actually false
- to claim that there is no Difference (accept H_0) when there really is no Difference

Research Tip

1. get curvas at .05 α
2. Do not invest resources until you reach .01 α
3. then, only after replication!

Error Types .01 - .01

STANDARD SCORES

Z Scores

- Z ^{scores} (lower case) are synonymous with the term "standing score".
- Standing scores are given in terms of standing deviation units (remember the concept that "Life is a bell-shaped curve").

$$Z = \frac{(X - \bar{X})}{SD} \quad \text{or} \quad \frac{X}{SD}$$

Z scores are usually expressed from -3.0 to +3.0.

A Z score is the deviation of any score from the mean, expressed in terms of standing deviation units.

Properties of the Z score:

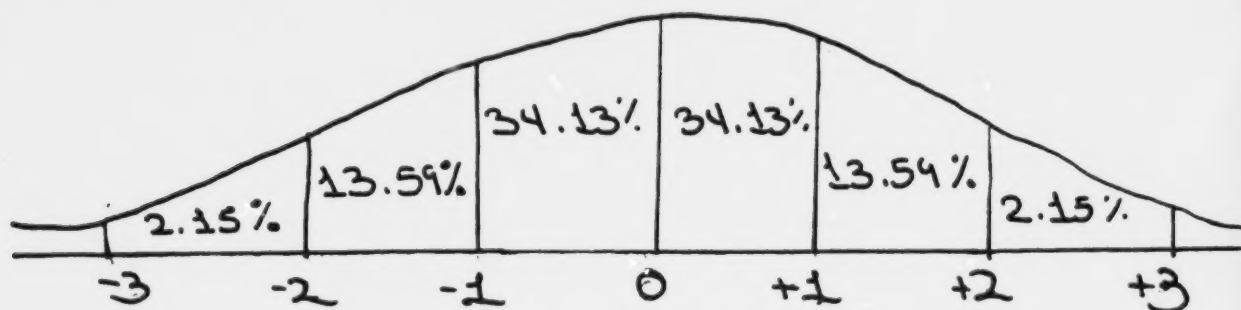
1. $\sum Z \text{ scores} = 0$
2. $\bar{X} Z \text{ scores} = 0$
3. The variance and SD of Z scores for any set of data = 1

The bell-shaped curve and Z scores are central to an understanding of normal distribution and cognate principles associated with probability.

Note. Normal distribution is also referred to as GAUSSIAN Distribution.

Z. scores. 01-01

A representation of the bell-shaped curve (i.e., normal distribution), with 2 scores and SD units



Are you acquainted with standard tests such as the GRE or SAT? Often:

$$\bar{X} = 500$$

$$SD = 100$$

Thus, a GRE score of 600 means that the subject scored better than 84% of all other people

Imagine a standard IQ test:

$$\bar{X} = 100$$

$$SD = 16$$

If a subject had a raw score of 129, what does this mean?

$$z = \frac{(X - \bar{X})}{SD} = \frac{129 - 100}{16} = \frac{29}{16} = 1.81$$

Recall, $z = \frac{\text{Deviation of score}}{\text{Standard deviation}}$

Then, "convert" the z score so that you can identify the percent of elements between the z score and the mean of the set of scores.

1. Use an "Areas of a normal curve" table

In this example ($z = 1.81$), the table Datum = .4649

2. Multiply the table datum from "areas of a normal curve" by 100

$$.4649 \times 100 = 46.49 \%$$

3. Thus, the converted percentage (46.49%) represents the percentage of cases between the mean and the z value.

4. Determine percentile rank

~~50 mean~~

$$50 + 46.49 = 96.49 \% \text{-tile}$$

Thus, the subject with a raw score of 129 had a "better" score than 96.49% of all participants

Non-Parametric Tests

Chi-Square Test of Significance

- ~ for two categories and one variable
- ~ based on the discrepancy between observed frequencies and expected frequencies (based on the stated hypothesis)

H_0 : There is no Difference among CCIS/CBL students in their preference for 1200 band and 2400 band modems.

- ~ the categories are mutually exclusive and exhaustive

Population: All CCIS/CBL students

Sample: 100 students selected by random selection from the entire population

$$\chi^2 = \sum \frac{(F_o - F_e)^2}{F_e} = \sum \frac{(D^2)}{E}$$

f_o : observed frequency (O)

f_e : expected frequency (E)

Data: + 1200 band 2400 band

χ^2 01-01	f_e	50	50
	f_o	20	80
	Difference (D)	-30	30
	D^2	900	900
	D^2/f_e	18	18

$$\chi^2 = \sum \frac{D^2}{f_e} = \sum \frac{(900 + 900)}{50} = \frac{1,800}{50}$$

$$\chi^2 = 36.00$$

$$df = 2 - 1 = 1$$

criticism χ^2 for 1 df 01 = 2.706
 .05 = 3.841

At .05 level of significance:

computed $\chi^2 = 36.00$ is greater than

criticism $\chi^2 = 3.84$ ($\alpha = .05$)

$\therefore$ there is a significant difference.

That is to say, Differences are significant and not due to chance.

$\therefore$ reject H_0 (the Null Hypothesis) and accept the Alternate hypothesis.

p.01-02

Chi-Square Test of Significance

~ for two variables, with two categories for each variable

H_0 : There is no Difference among DIS and non-DIS CCIS/CBL students in their preference for 1200 Samd and 2400 band modems.

Population: All CCIS/CBL Students

Sample: 200 DIS students and 250 non-DIS students selected by random from the entire population

Data:

Program	Preference		Total
	1200 band	2400 band	
DIS	80	120	200
non-DIS	140	110	250
Column Total	220	230	450

In this example (and throughout later χ^2 analysis) a major task is the determination of expected frequency. Unlike problems where the selection is discrete (50, 50), samples are usually divided on a proportional basis.

χ^2 02-01 To effect the task of determining f_e , it is typical to use the process identified below:

	1200 band	2400 band	Bondage Total
DIS	80 a	120 b	200
non-DIS	140 c	110 d	250
Bondage Total	220	230	450

To determine f_e for each cell multiply the bondage totals by the cell total and divide by N common to a cell

$$a = \frac{200 \times 220}{450} = 97.7778$$

$$b = \frac{200 \times 230}{450} = 102.22$$

$$c = \frac{250 \times 220}{450} = 122.22$$

$$d = \frac{250 \times 230}{450} = 127.7778$$

7.02-02

		1200 band		2400 band		Borgers Total
DIS	O	80	2	120	b	200
	E	97.78		102.22		
	D	17.78		- 17.78		
	D ²	316.13		316.13		
	D ² /E	3.23		3.09		
non-DIS	O	140	c	110	d	250
	E	122.22		127.78		
	D	- 17.78		17.78		
	D ²	316.13		316.13		
	D ² /E	2.59		2.47		
Borgers Total		220		230		

$$\chi^2 = \sum \frac{D^2}{E} = 3.23 + 3.09 + 2.59 + 2.47$$

$$= 11.38$$

$$df = (R-1)(C-1)$$

$$= (2-1)(2-1)$$

$$= 1 \times 1$$

$$= 1$$

R = Rows
C = Columns

72.02.073

At .05 level of significance:

computed $\chi^2 = 11.38$

critical $\chi^2 = 3.84$ ($\alpha = .05$)

$\therefore$ there is a significant difference

That is to say, there is a significant and not due to chance.

To be more exact:

60.00% of all DIS CCIS/CBL students prefer 2400 band

44.00% of all non-DIS CCIS/CBL students prefer 2400 band

on - The majority of all DIS students prefer 2400 band, whereas the majority of all non-DIS students prefer 1200 band, a lower band rate.

$\therefore$ Reject H_0 ; Accept H_A

Outcome: DIS students prefer higher
+ higher band rates as compared
- to non-DIS students

Chi-Square is often identified as the "goodness of fit" test.

- to determine if there is a significant difference between observed and expected frequencies
- if chi-square is significant (computed chi-square > criterion chi-square)

A few "tips" about chi-square:

1. Expected frequencies should be less than five ($n=5$ for any one cell) for only 20 percent of the cells
2. The frequency categories must be mutually exclusive. That is to say, a score can be counted only once.

H₀: There is no difference in preferred teaching/learning method (of four possibilities) when evaluated by CCIS/CISL students immediately after instruction and after the end of the 6-month registration period.

Learn to "scope out" the situation

1. Dependent variable ... preference for teaching/learning method

A. Among four methods

B. During instruction

C. After Instruction

2. Independent Variable ...

A. Four methods

2.03-01

Preference During Instruction						Total
		A	B	C	D	
A	O	10	6	7	8	31
	E	6.20	8.27	6.20	10.33	
	D ²	3.80	2.27	0.80	2.33	
	D ²	14.44	5.15	0.64	5.43	
	D ² /E	2.33	0.62	0.10	0.53	
B		5	20	5	7	37
C		8	8	10	15	41
D		7	6	8	20	
Total		30	40	30	50	150

Note. To calculate observed frequencies (f_o , O)

A. Multiply the column total by the row total

B. Divide by the total N.

In the above table:

Row A x Column A ($n=10$, $O=10$)

$$E \times = \frac{30 \times 31}{150} = 6.20$$

7.03.02

Calculate for each Cell

$$\begin{array}{l} E \\ D \\ D^2 \\ D^2/E \end{array}$$

$$\text{Computed } \chi^2 = 24.49$$

$$\begin{aligned} df &= (R-1)(C-1) = (4-1)(4-1) \\ &= 3 \cdot 3 \\ &= 9 \end{aligned}$$

For 9 df & $\alpha = .05$, criterion $\chi^2 = 16.92$

$$\text{Computed } \chi^2 (24.49) > \text{Criterion } \chi^2 (16.92)$$

$\therefore$ Difference is significant & not due to chance
that is to say, there is a significant Difference
Reject the null hypothesis

A. During class, there is a preference for method D ($n = 50$)

B. After class, preference seems to shift to method C (pre $C_n = 30$, post $C_n = 41$)

Note. There are other statistical tests that can be used to "better" explain phenomena. But, this example clearly shows & how χ^2 can be used for more than 2×2 factorial Designs.

7.03-03

Assignment of Ranks

~ ordinal data

~ Difference between elements is not measured, instead they are placed in rank order

Subject	"Score"	Rank
8	4	1
5	7	2
1	10	3
4	11	4.5
7	11	4.5
9	12	6
2	13	8
3	13	8
6	13	8

For Subjects 4 & 7, each has a "score" of 11, $\therefore$ it is not possible to assign "4" & "5" as ranks.

$$4 + 5 = 9 / 2 = 4.5$$

For subjects 2, 3, & 6

$$8 + 8 + 8 = 24 / 3 = 8.0$$

Ranks.01-01

Spearman Rho

- ~ a non parametric test for correlation
- ~ Data are ranked
- ~ each subject will have two ranked scores

$$\text{Spearman Rho} = r_{\text{rho}}$$

$$r_{\text{rho}} = 1 - \frac{(6)(\sum D^2)}{n(n^2 - 1)}$$

Subject	Rank #1	Rank #2	D	D ²
4	1	4	-3	9
9	2	2	0	0
15	3	9	-6	36
1	4	1	3	9
1	5	7	-2	4
1	6	10	-4	16
1	7	8	-1	1
1	8	13	-5	25
1	9	5	4	16
1	10	3	7	49
1	11	11	0	0
1	12	6	6	36
2	13	12	1	1
6	14	15	-1	1
10	15	14	1	1
				204

$$r_{\text{rho}} = 1 - \frac{(6)(204)}{15(225-1)} = 1 - \frac{1,224}{3,360}$$

Spearman. 01-01

$$r_{rho} = 1 - 0.364$$

$$r_{rho} = .636$$

$$df = n - 1 = 15 - 1 = 14$$

Computed rho (.636) > Critical rho (.544, $\alpha = .05$)

$\therefore$ The difference is significant $\therefore H_0$ is rejected

That is to say, there is a positive significant relationship

H_0 There is no difference among CCIS/CBL MS students at graduation regarding the correlation between portfolio rating (Rank #1) and terminal GPA (Rank #2)

Conclusion: With a fair degree of confidence, you can assume that there is an association between portfolio rating & terminal GPA.

✓ yet - do not infer cause and effect.

A high portfolio rating ~~does~~ does not make a student have a high terminal GPA.

Sign Test

- ~ to determine the level of difference between two categories
- ~ the same subject is used for each measure
- ~ Determines if a difference exists within a group of subjects, relative to the ratings given each subject for two variables

Subject	Pre-test	Post-Test	Change
1	2	3	+
2	1	2	+
3	1	3	+
4	5	4	-
5	5	4	-
6	6	7	+
7	1	2	+
8	2	7	+
9	3	6	+
10	2	4	+
11	2	5	+
12	6	6	0
13	7	1	-
14	1	4	+
15	2	8	+
16	4	10	+
17	6	9	+
18	2	5	+
19	3	5	+
20	9	2	-
21	5	8	+
22	3	6	+
23	4	2	-
24	2	7	+
25	7	2	-
26	1	4	+
27	2	4	+

sign-test.01-02

Scenario: CCIS/CBL applicants were observed during Unix training workshops. A team of 5 faculty members rated each applicant to "adaptability to learn" Unix.

At the completion of the first day, subjects were given the opportunity to talk with currently enrolled students about their experience with Unix.

Subjects were then taken back to the workshop and their "adaptability to learn" Unix was again measured.

The first workshop is regarded as the pretest.

Talking with students is regarded as the treatment.

The second workshop is regarded as the posttest.

H₀ Talking with present students will have no effect on the "adaptability to learn" Unix among CCIS/CBL applicants.

PROCESS:

1. $n + \text{changes} = 21$

2. $n - \text{changes} = 06$

3. $n \text{ "no change"} = 1$

4. $R = 6$ (the smaller number)

5. Refer to a sign test table for critical values of R

sign.test.01-02

$N = 27$ ($n = 28 - 1$ "0" change)

Criterion R for:

$$\left. \begin{array}{l} \alpha = .05 \\ N = 27 \end{array} \right\} R = 7$$

e. Computed $R (= 6) < \text{Criterion } R (= 7)$

$\therefore$ there is a significance

(As opposed to many other tests for a sign test ... significance is declared where computed $R < \text{Criterion } R$).

Conclusion:

1. Talking with present CCIS/CBL students during Unix workshops improves the "adaptability to learn Unix" among applicants.

Note. A major point to remember for the sign test is that significance is declared when computed R is less than Criterion R .

sign. test. 01-03

Wilcoxon Matched Pairs

- ~ often used in place of student's t-test when non-parametric analysis is more appropriate
- ~ The test statistic is called T (as opposed to t)
- ~ Wilcoxon Matched Pairs analysis is used with Data form:
 1. matched pairs (two sets of subjects)
 2. one subject in a pre/post test situation
- ~ Wilcoxon Matched Pairs is typically regarded as being more robust than the signed ranks test

Note. It is difficult to effect a study with matched pairs. Many things happen resulting in the "un-matching" of the pairs. There are also many intervening variables that may have unnoticed, but true, effect on outcomes.

Scenario: This project involves colleagues who join a Doctoral program in Information Science.

1. One colleague joins a traditional (on campus, all night time classes) doctoral program.
2. The other colleague joins Dana's CCIS/CBL program.
3. The decision to join the traditional or external program was based on a coin toss. Therefore, self-selection for delivery is not an issue.

wilcoxon.01-02

In this study, colleagues are defined as teachers in a MIS university program.

H₀: There is no difference in "ability to learn independently" between colleagues^{MIS} teacher who attended a traditional Information Systems Doctoral program and MIS teachers who attended Nova University's external doctoral program in the CCIS/CBL.

A reliability coded test of "ability" to learn independently was given to all subjects at the end of the spring year in the doctoral program.

PAIR	Rank: Nova Student	Rank: Traditional	
1	72	53	+19
2	59	32	+27
3	58	59	-01
4	14	10	+04
5	51	38	+13
6	57	50	+07
7	72	76	-04
8	66	63	+03

In this study, colleagues are defined as teachers from the same school. Both colleagues teach MIS and are basically equal in all identifiable characteristics associated with independent learning.

Pair	Rank A	Rank B	Difference	Diff. Rank	Sign
1	72	53	+19	7	+
2	59	32	+27	8	+
3	58	59	-01	1	-
4	14	10	+06	4	+
5	51	38	+13	6	+
6	57	50	+07	5	+
7	72	76	-04	3	-
8	66	63	+03	2	+
	451	392			

Note. Difference in Rank is based on absolute values.

Process:

1. Add the ranks of the less frequent sign. This is the T value.

The less frequent sign is -.

The sum of the - signs is:

$$T = 1 + 3 = 4$$

2. Refer to a table of T values. In this example:

$$\alpha = .05, n = 8 \text{ pairs}$$

$$\text{Criterion } T = 4$$

3. Compute $T(4) = \text{Criterion } T(4)$

For Wilcoxon Matched Pairs, significance is declared if the computed T is equal to or less than the criterion T .

Wilcoxon.01-03

The difference is significant and not due to chance.

Conclusion: At the end of the second year in a Doctoral program, it is observed that students in an external degree program are more ADEPT at independent learning than their colleagues in a traditional Doctoral program.

Mann-Whitney U

- ~ a non-parametric test "similar" to a t-test
- ~ does not require matching pairs
- ~ used in place of the t-test when data are ranked and from two independent samples

Samples can be from:

1. The same population, with an experimental & control situation
 2. The same population with results compared to two categories of one variable
 3. From two populations if the same variable is purposely compared to two populations
- ~ Mann-Whitney U is especially used for low sample size

Scenario: A researcher wanted to study the effect of lecture and CBT on learning outcomes in Unix Training.

From the population of 124 potential subjects (CCIS/CBL prospective students), a sample of 28 subjects was selected through random processes

~~Subjects were further divided (again, through random processes) to two groups of 10 each.~~

Basic skills in Unix was measured to serve as baseline data. After a 2 Day workshop, subjects were measured again.

Mann-Whitney U

The only difference in effect was that Group A subjects received lecture, whereas Group B received CBT "hands-on" Unix training.

Rank Order	Lecture ^A	CBT ^B
1	1	
2	2	
3		3
4.5		4.5
4.5	4.5	
6		6
7	7	
8		8
9	9	
11		11
11		11
11	11	
13.5		13.5
13.5		13.5
15		15
	$\Sigma = 34.5$	$\Sigma = 85.5$
	$n = 6$	$n = 9$

"Very" able to learn Unix (highest test score) was Rank "1" and "Least" able to learn Unix was ranking "15"

Formulas:

$$U_2 = (n_2)(n_b) + \frac{n_b(n_b+1)}{2} - \Sigma R_b$$

$$U_b = (n_a)(n_b) + \frac{n_a(n_a+1)}{2} - \Sigma R_a$$

ΣR_2 = sum of the ranks for sample A

ΣR_3 = " " B

n_2 = number in sample A

n_3 = " " B

Note. n_2 does not have to equal n_3

$$U_2 = (6)(9) + \frac{(9)(9+1)}{2} - 85.5$$

$$= 54 + \frac{90}{2} - 85.5$$

$$= 13.5$$

$$U_3 = (6)(9) + \frac{(6)(6+1)}{2} - 34.5$$

$$= 54 + \frac{42}{2} - 34.5$$

$$= 40.5$$

Verify work:

$$(n_2)(n_3) = U_2 + U_3$$

$$(6)(9) = ~~85~~ 13.5 + 40.5$$

mann-whitney.01-03

Determine the criterion value and compare this value to the smaller U value.

$$9 \times 6 \text{ Criterion } U (d.05) = 10$$

$$\text{Computed } U (\text{smaller value}) = 13.50$$

$$\text{Computed } U (13.50) > \text{Criterion } U (10)$$

$\therefore$ There is no significant difference

There is no difference in effect of teaching methodology (lecture, CST) on learning outcomes in Univ training.

Kruskal-Wallis

- a non-parametric test for use with ordinal data
- using ranked data, Kruskal-Wallis is used to test for the significance of differences among three or more groups
- a major feature of Kruskal-Wallis is that it is used when there are three or more groups (or sub-samples)
- Kruskal-Wallis can be used:
 1. When there is one variable with three (or more) categories (denoting three or more sub-sample groups)
 2. When there are three different populations
 3. When there are three (or more) subgroups of one population, with only one variable of only one category
- Data comparison is "between groups" and not "within groups"

Scenario: It is not known if CCIS/CBL MS applicants are in any way affected (regarding program success) by the type of institution where they earned their BA/BS.

It is generally perceived that there are three types of institutions:

1. Research-oriented universities, where BA/BS graduates have reduced contact with senior faculty

B. Teaching-orienting institutions, where BA/BS graduates have extensive contact with senior faculty

C. Small liberal arts colleges, where student/faculty contact is extensive throughout the full-year BA/BS program.

H₀: There is no difference in first-year GPA among CCIS/CBL MS students regarding prior college experience:

- A. graduates of a Research-Oriented university
- B. " " teaching-orienting "
- C. " " liberal arts college

GPA's were calculated for a sample of 24 CCIS/CBL 1st year students. Subjects were ranked from 1 to 24 and subjects were then grouped into respective categories.

	Research	Teaching	Liberal Arts
Kruskal-Wallis, U = 61.02	3.5	1.0	3.5
	8.0	2.0	5.0
	9.0	11.0	6.5
	13.5	12.0	6.5
	16.5	13.5	10.0
	20.0	16.5	15.0
	21.5	19.0	18.0
		21.5	23.0
			24.0

The test statistic for Kruskal-Wallis is H

$$H = \left[\frac{12}{N(N+1)} \right] \left[\sum \frac{(\sum T_A)^2}{n_A} \right] - 3(N+1)$$

$\sum T_A$ = the sum of ranks for each group

$$\sum T_A = T_{A1} + T_{A2} + T_{A3}$$

n = the number of subjects for each T_A

N = Total number of subjects

In this example:

$$\sum T_{A1} = 92.0 \quad n_1 = 7$$

$$\sum T_{A2} = 96.5 \quad n_2 = 8$$

$$\sum T_{A3} = 111.5 \quad n_3 = 9$$

$$\sum T_A = \quad N = 24$$

If, it is best - break the formula into manageable portions

$$\left[\frac{12}{(N)(N+1)} \right] = \frac{12}{24(24+1)} = \frac{12}{24 \times 25} = \frac{12}{600}$$
$$= 0.02$$

$$\left[\frac{(\sum TA)^2}{n_A} \right] = \left[\frac{(\sum TA_1)^2}{n_{A_1}} + \frac{(\sum TA_2)^2}{n_{A_2}} + \frac{(\sum TA_3)^2}{n_{A_3}} \right]$$

$$= \left[\frac{92^2}{7} + \frac{96.5^2}{8} + \frac{111.5^2}{9} \right]$$

$$= \left[\frac{8464}{7} + \frac{9312.25}{8} + \frac{12432.25}{9} \right]$$

$$= 1209.14 + 1164.03 + 1381.36$$

$$= 3754.53$$

Finally,

$$3(N+1) = 3(24+1) = 3(25)$$

$$\therefore = 75$$

$$H = (.02)(3754.53) - 75$$

$$H = 75.09 - 75$$

$$H = 0.09$$

$$df = A - 1$$

$$df = 3 - 1$$

$$df = 2$$

Criticism χ^2 values are used for H

In this example:

$$H(.05, df=2) = 5.99$$

$$\text{Computed } H(0.09) < \text{Criticism } H(5.99)$$

$\therefore$ No difference

Difference is due to chance; not due to
real effects (i.e., type of BA/BS program).

Type of BA/BS degree-granting college/
university has no effect on 1st year
CCIS/CBL GPA.

Rank Sums

- ~ a non-parametric test with grouped data
- ~ used to test for difference in ranks between two categories (or levels) of the independent variable
- ~ Rank Sums is applicable to two groups of subjects:
 1. Control ; Experimental
 2. Two sets of data

Scenario: The population is DIS students in the CCIS/CBL. Using a random selection process, 21 DIS students were selected to serve as a sample.

During SI'90 two teaching methods were used for the 21 subjects

1. 11 subjects were placed in a CBT-type class (CBT)
2. 11 subjects were placed in a lecture-type class. (Lecture)

After instruction, subjects were required to complete an evaluation of their respective class.

The dependent variable is student satisfaction with their class.

Following student evaluation, results were ranked:

1 : subjects most satisfied

n. 1 1

H₀: There is no difference among CCIS/CBL & DIS students regarding satisfaction with CBT and Lecture teaching methodology.

Rank	CBT	Lecture	CBT Rank	Lecture Rank
1	✓		1	
2	✓		2	
3.5	✓		3.5	
3.5	✓		3.5	
5		✓		5
6	✓		6	
7	✓		7	
8	✓		8	
9		✓		9
10	✓		10	
11.5	✓		11.5	
11.5	✓		11.5	
13		✓		13
14		✓		14
15		✓		15
16		✓		16
17		✓		17
18	✓		18	
19		✓		19
20		✓		20
21		✓		21

$$\Sigma = \quad \underline{82 (T_1)} \quad \underline{149 (T_2)}$$

$$n = \quad 11 (n_1) \quad 10 (n_2)$$

Rank-sum.01-02

$$Z = \frac{2T_{\text{smaller}} - n_1(N+1)}{\sqrt{\frac{(n_1)(n_2)(N+1)}{3}}}$$

T_{smaller} = the sum of ranks for the group with the smallest n (if $n_1 \neq n_2$)

n_1 = the number of subjects for T_{smaller}

$$Z = \frac{2(82) - 11(22)}{\sqrt{\frac{(10)(11)(22)}{3}}} = \frac{164 - 242}{\sqrt{\frac{2420}{3}}}$$

$$Z = \frac{-78}{\sqrt{806.67}} = \frac{-78}{28.4} = -2.78$$

In contrast to most statistical tests, there is no table of Z values for Rank Sums analysis. Instead:

$$Z = + \text{ or } - 1.96 (\alpha .05)$$

Any score "between" -1.96 & $+1.96$ is not significant.

Any score greater than $+1.96$ or less than -1.96 is significant.

In this example,

$$Z = -2.78 < Z = -1.96 (\alpha = .05)$$

$\therefore$ There is a significant difference

The group receiving CBT ($N = 82$ where
 $1 = \text{High Rating}$) was significantly more
satisfied with their class than similar
subjects who received lecture ($N = 149$).

Note. Perhaps a t-test could have been used, but
it is assumed in this example that the
precision of measurement did not support
parametric analysis.

Parametric Tests

Correlation

- ~ a measure of the degree of association between variables
- ~ regression involves the study of interrelationship (association) of two or more variables

Pearson: parametric ... interval/ratio data

Spearman: non-parametric ... ranked data

Pearson's r can be determined through various formulas

Correlation values range from -1.00 to $+1.00$.

The $+$ and $-$ signs are only used to show direction.

r values are expressed as:

$+ .01$	$-.53$
$+ .98$	$-.69$
$+ .73$	$-.09$

Typically, the $+$ sign is not expressed.

But, a Zero does not come before the decimal point.

Many formulas can be used to determine the coefficient of correlation (r).

Z score Data

$$r = \frac{\sum (Z_x Z_y)}{n}$$

n = the number of pairs of scores

Raw Data

$$r = \frac{(n)(\sum xy) - (\sum x)(\sum y)}{\sqrt{[(n)(\sum x^2) - (\sum x)^2][n)(\sum y^2) - (\sum y)^2]}}$$

Subject	x	x ²	y	y ²	xy
1	37	1,369	61	3,721	2,257
2	40	1,600	40	1,600	1,600
3	36	1,296	40	1,600	1,440
4	49	2,401	38	1,444	1,862
5	36	1,296	37	1,369	1,332
6	40	1,600	50	2,500	2,000
7	39	1,521	24	576	936
8	47	2,209	30	900	1,410
9	32	1,024	16	256	512
10	65	4,225	50	2,500	3,250
11	27	729	26	676	702

n = 11 $\sum x = 448$ $\sum x^2 = 19,270$ $\sum y = 412$ $\sum y^2 = 17,142$ $\sum xy = 17,301$

$\bar{x} = 40.73$ $\bar{y} = 37.45$

correlation Pearson .05.00

$$r = \frac{(11)(17,301) - (448)(412)}{\sqrt{[(11)(19,270) - (448)^2][(11)(17,142) - (412)^2]}}$$

$$r = \frac{190,311 - 184,576}{\sqrt{[211,970 - 200,704][188,562 - 169,744]}}$$

$$r = \frac{5,735}{\sqrt{[11,266][18,818]}}$$

$$r = \frac{3,733}{\sqrt{212,003,588}}$$

$$r = \frac{5,735}{14,560.34}$$

$$r = .39$$

correlation. Pearson. 01.03

$$r = .39$$

$$df = n - 2 = 11 - 2 = 9$$

$$\text{Criterion } r (\alpha .05, df = 9) = .602$$

$$\text{Computed } r = .39$$

$$\text{Criterion } r (.602) > \text{computed } r (.39)$$

$\therefore$ There is no significant correlation

Note. Criterion r values are very sensitive to n .

$$\text{Criterion } r (\alpha .05, df = 10) = .576$$

$$\text{Criterion } r (\alpha .05, df = 100) = .195$$

"Rule of Thumb"

1. $r = .00$ to $.03$... no correlation
2. $r = .30$ to $.60$... mild correlation
3. $r = .60$ to $.90$... strong correlation
4. $r > .90$... high degree of correlation

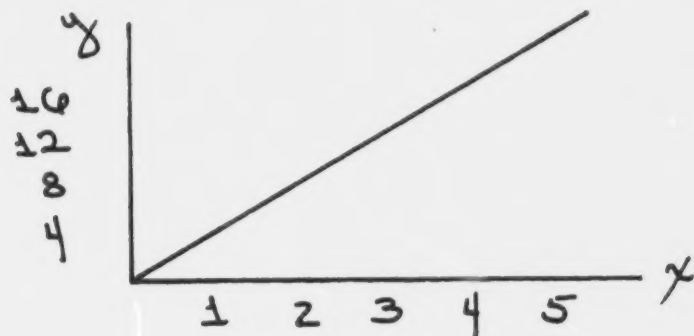
Correlation and Regression

~ with sufficient data, it is possible to predict (with varying degrees of assurance) the value of one score ... based on the value of another score

~ that is to say,

given x it is possible to predict the value of y

~ the prediction is possible due to the construction of a regression from the set of data



In this figure you have an example of $r = +1.00$

given $x = 2$, you can predict with confidence that $y = 8$

Regression Line:

1. A "tool" for prediction
2. Based on a significant r

Note. A significant r means that the association of two variables is not due to chance. But, recall that $r \geq .70$ the "norm" for professional acceptance (any investment) of the association.

$$y' = \bar{y} + \left[(r) \left(\frac{SD_y}{SD_x} \right) (x - \bar{x}) \right]$$

SUBJECT	X	Y
1	12	20
2	8	15
3	9	17
4	20	36
5	18	31
6	7	12
7	10	19
8	7	13
9	5	9
10	21	40
11	11	19
12	10	15
13	17	40
14	12	20
15	12	21
16	14	25

$$\bar{X} = 12.06$$

$$\bar{Y} = 22.00$$

$$r = .96$$

$$X_{sy} = 4.78$$

$$Y_{sy} = 9.77$$

$$n = 16$$

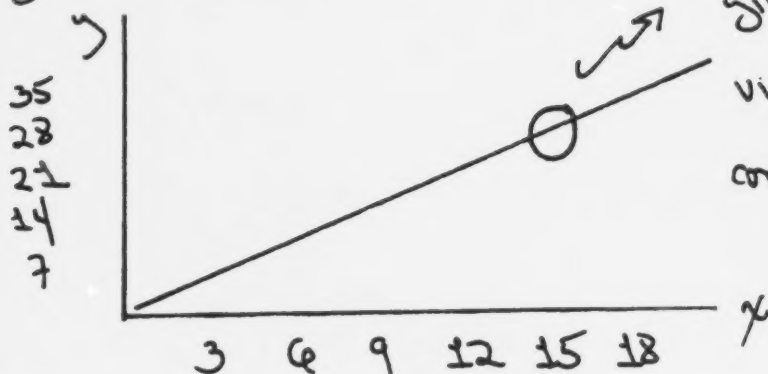
Solve for y' based on a score of $X = 15$

$$y' = 22.00 + \left[(.96) \left(\frac{9.77}{4.78} \right) (15 - 12.06) \right]$$

$$y' = 22.00 + \left[(.96) (2.04) (2.94) \right]$$

$$y' = 22.00 + 5.76$$

$$y' = 27.76$$



given $X = 15$,
visual representation
confirms $y' = 27.76$

Note. Approximate the slope & plot the numbers to visually check the accuracy of your calculation

correlation, regression.01.03

Student's t-Test

- ~ interval or ratio data
- ~ the t-test is used to compare the means of two samples
- ~ the comparison of sample means is done to determine if the difference is due to true differences, or due to chance

$$\bar{X} \neq \bar{Y}$$

1. Due to chance?

2. Due to true difference?

Potential applications of a t-test

1. An extension of correlation studies
2. To compare data from an experimental group to cognate data from a control group

A major disadvantage of the t-test is that its use is restricted to only two sets of data.

To compare / understand:

$$X \text{ : } Y$$

- t-test OK

$$X, Y \text{ : } Z$$

- cannot use t-test

The use of t-test involves use of the standard error of the mean:

Standard Error of the Mean: $SE_{\bar{x}}$

$SE_{\bar{x}_1 - \bar{x}_2}$ Standard error of the difference between the means of x_1 & x_2 .

Standard Error of the Mean can be determined by various methods. The method often depends on other calculations that have been already completed.

Note. The use of N or $N-1$ has already been addressed. Since t-tests are typically done with $N < 30$, it is typical to use $N-1$.

If SD is already known:

$$SE_{\bar{x}} = \frac{SD}{\sqrt{N-1}}$$

From raw data:

$$SE_{\bar{x}} = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N(N-1)}}$$

But, in most cases, the standard error of the difference between two sample means is used with t-test calculations. And, most t-tests are done with raw data

$$t = \frac{\bar{X}_1 - \bar{X}_2}{SE_{\bar{X}_1 - \bar{X}_2}}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{\sum X_1^2 - \frac{(\sum X_1)^2}{N_1} + \sum X_2^2 - \frac{(\sum X_2)^2}{N_2}}{N_1 + N_2 - 2} \right] \left[\frac{N_1 + N_2}{(N_1)(N_2)} \right]}}$$

X_1 = scores in sample one (X)

X_2 = " " " two (Y)

$\sum X^2$ = sum of squares of each X value

$\sum Y^2$ = Y value

Subject	X_1	X_1^2	X_2	X_2^2	$X_1 X_2$
1	12	144	20	400	240
2	8	64	15	225	120
3	9	81	17	289	153
4	20	400	36	1,296	720
5	18	324	31	961	558
6	7	49	12	144	84
7	10	100	19	361	190
8	7	49	13	169	91
9	5	25	9	81	45
10	21	441	40	1,600	840
11	11	121	19	361	209
12	10	100	15	225	150
13	17	289	40	1,600	680
14	12	144	20	400	240
15	12	144	21	441	252
16	14	196	25	625	350
Σ	193	2,671	352	9,178	4,922
Mean:	12.06	166.94	22.00	573.63	307.63

L. test. 01-04

$$t = \frac{12.06 - 22.00}{\sqrt{\left[\frac{2,671 - \frac{(193)^2}{16} + 9,178 - \frac{(352)^2}{16}}{16 + 16 - 2} \right] \left[\frac{16 + 16}{(16)(16)} \right]}}$$

$$t = \frac{-09.94}{\sqrt{\left[\frac{2,671 - 2,328.06 + 9,178 - 7,744}{30} \right] \left[\frac{32}{256} \right]}}$$

$$t = \frac{-09.94}{\sqrt{\left[\frac{1,776.94}{30} \right] \left[0.125 \right]}}$$

$$t = \frac{-09.94}{\sqrt{(59.23)(0.125)}}$$

t-test.01-05

$$t = \frac{-09.94}{\sqrt{7.40}}$$

$$t = \frac{-09.94}{2.72}$$

$$t = -3.65$$

$$df = N_1 + N_2 - 2$$

$$df = 16 + 16 - 2$$

$$df = 32 - 2$$

$$df = 30$$

$$\text{Criterion } t (d.05, df=30) = \pm 2.04$$

$$\text{Criterion } t (d.01, df=30) = \pm 2.75$$

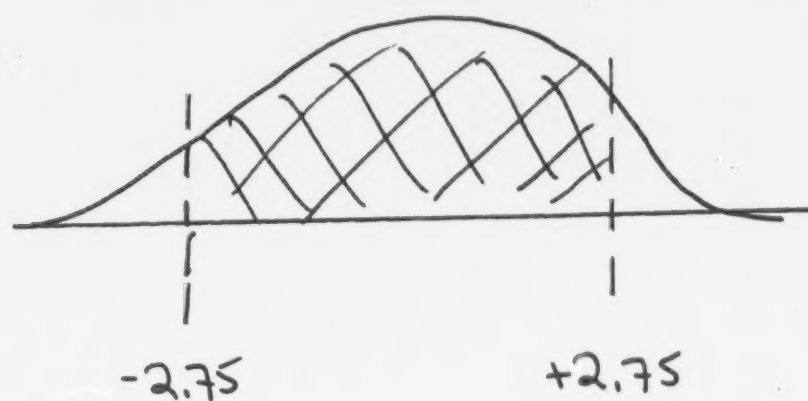
$$\text{Criterion } t (d.01, df=30) = 2.75$$

is less than Computer ~~t~~ (i.e. -3.65)

∴ Differences are significant as

not due to chance. 68

t-test. 01-06



 = Area of no difference

-3.75 goes beyond the area of no difference.

Accordingly, differences are real.

The data are from two separate populations, even though there is a degree of association between the two samples.

AGAIN:
$$t = \frac{\bar{X}_1 - \bar{X}_2}{SE_{\bar{X}_1 - \bar{X}_2}}$$

69

Raw Data were used to demonstrate what computer programs can do in seconds.

Additional Uses of Student's t-Test

~ perhaps the most common use of the t-test is its use (overuse?) in the comparison of scores between two groups

A. Control : Experimental (two separate groups)

B. Pretest : Posttest (same group)

with two separate groups:

$$df = N_1 + N_2 - 2$$

With one group:

$$df = n - 1$$

Note:

1. The t-test is only used in situations where a null hypothesis has been declared.

2. Do not confuse the use of 1-tailed and 2-tailed t-tests.

H_0 : "No Difference" ... Two-tailed

H_0 : "Directional Difference" ... one-tailed

Remember - t-test depends on Random Selection.

t-test. 02-01

The t-test is also closely associated with correlation

Subject	X	X ²	y	y ²	xy
1	2	4	5	25	10
2	3	9	6	36	18
3	5	25	11	121	55
4	1	1	3	9	3
5	7	49	13	169	91

$$n = 5$$

$$\sum \quad \quad 18 \quad \quad 88 \quad \quad 38 \quad \quad 360 \quad \quad 177$$

$$r = \frac{(n)(\sum xy) - (\sum X)(\sum y)}{\sqrt{[(n)(\sum X^2) - (\sum X)^2][n)(\sum y^2) - (\sum y)^2]}}$$

$$r = \frac{(5)(177) - (18)(38)}{\sqrt{[(5)(88) - (18)^2][(5)(360) - (38)^2]}}$$

$$r = \frac{885 - 684}{\sqrt{[(440)(324)][(1800) - (1444)]}}$$

$$r = \frac{201}{\sqrt{(116)(356)}}$$

t-test.02-02

$$r = \frac{201}{\sqrt{41,296}} = \frac{201}{203.214} = .989$$

$$r = .989$$

$$df \text{ for } r = n - 2 = 5 - 2 = 3$$

$$\text{Criterion } r (\alpha .05, 3df) = .878$$

$$t = \frac{(r)(\sqrt{n-2})}{\sqrt{1-r^2}} = \frac{(.989)(\sqrt{5-2})}{\sqrt{1-.978}}$$

$$t = \frac{(.989)(\sqrt{3})}{\sqrt{.022}} = \frac{(.989)(1.732)}{(0.148)}$$

$$t = \frac{1.713}{0.148} = 11.57$$

$$df \text{ for } t = n - 2 = 5 - 2 = 3$$

$$\text{Criterion } t (\alpha .05, 3df) = 3.18$$

$$\text{Computed } t (11.57) > \text{Criterion } t (3.18)$$

72 $\therefore$ Differences are significant; not Due to chance.

t-test.05-.03

Analysis of Variance: Simple Ratio

- ~ ANOVA and the F test are now widely used in place of the t-test
- ~ A major advantage of ANOVA is its ability to be used to test on two or more groups at the same time
- ~ with ANOVA, it is possible to determine:
 1. Variance "within" a sample
 2. Variance "between" samples
- ~ a parametric test, ANOVA depends on interval or ratio data
- ~ the simplest use of ANOVA (and the F test) is to determine the
 1. Difference between two samples
 2. variability between two samples
- ~ in this example, a ratio is developed between the variance of the two samples

$$F_{\text{RATIO}} = \frac{\text{MS Larger}}{\text{MS smaller}}$$

$$MS = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}$$

Sample	X_1	X_1^2	X_2	X_2^2	$X_1 X_2$
1	3.29	10.82	8.12	65.93	26.71
2	2.19	4.80	7.19	51.69	15.75
3	3.35	11.22	9.13	83.34	30.58
4	2.05	4.20	6.54	42.77	13.41
5	2.13	4.54	7.68	58.98	16.36
6	3.49	12.18	9.56	91.39	33.36
7	2.28	5.20	6.15	37.82	14.02
8	2.93	8.58	7.93	62.88	23.23
9	3.01	9.06	8.91	79.39	26.82
10	2.39	5.71	7.12	50.69	17.02
11	2.93	8.58	9.47	89.68	27.75
12	3.68	13.54	8.12	65.93	29.88
13	3.38	11.42	8.69	75.52	29.37
14	2.75	7.56	7.35	54.02	20.21
15	3.87	14.98	9.68	93.70	37.46
16	2.18	4.75	7.02	49.28	15.30
17	2.91	8.47	6.93	48.02	20.17
18	3.47	12.04	8.18	66.91	28.38
19	3.71	13.76	8.92	79.57	33.09
20	3.83	14.67	7.92	62.73	30.33

$n = 20$ $\Sigma = 59.82$ 160.61 489.20

186.08 $1,310.24$

Mean 2.99 9.30 8.03 65.51 24.46

$X_1 = BS/BA$ GPA

$X_2 = CCIS$ Portfolio Score (1, Low - 10, High)

ANOVA.01-02

$$MS = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1}$$

$$MS = \frac{(186.08) - \frac{(59.82)^2}{20}}{20-1}$$

$$MS = \frac{(186.08) - (3,578.43/20)}{19}$$

$$MS = \frac{(186.08) - (178.92)}{19} = \frac{7.16}{19}$$

~~$$MS = \frac{33,293.43}{19}$$~~

$$MS = 0.377$$

~~$$MS = 1,752.28$$~~

$$MS = \frac{\sum X_2^2 - \frac{(\sum X_2)^2}{n}}{n-1}$$

$$MS = \frac{(1,310.24) - \frac{(160.41)^2}{20}}{20-1}$$

$$MS = \frac{(1,310.24) - (25,795.57/20)}{19}$$

$$MS = \frac{(1,310.24) - (1,289.78)}{19}$$

$$MS = \frac{20.46}{19}$$

$$MS = 1.077$$

Anova.01-04

$$F = \frac{\text{MS Larger}}{\text{MS smaller}}$$

$$F = \frac{1.077}{0.377}$$

$$F = 2.857$$

$$df X_1 = n X_1 - 1 = 20 - 1 = 19$$

$$df X_2 = n X_2 - 1 = 20 - 1 = 19$$

$$\text{Criterion } F (\alpha .05, n df = 20, d df = 19) = 2.15$$

$$\text{Criterion } F (\alpha .01, n df = 20, d df = 19) = 3.00$$

$$\text{Computed } F (2.857) > \text{Criterion } F (\alpha .05) 2.15$$

$\therefore$ Data are significant

$$F = 2.857 *$$

Note. * Denotes significance $\alpha = .05$

* * " " " " .01

Analysis of Variance: One-Way

~ A more frequent (and useful) utility of the F test and ANOVA involves the use of these utilities to determine if:

1. The variance between/among samples is $\geq$ or $<$ the variance within each sample.

2. The means differ between/among samples.

~ Justification of the use of ANOVA analysis:

1. If "within sample" variance is large, the variance between/among samples tends to be of minimal importance.

2. Differences between sample means tends to result in significant differences between/among samples.

Possible paradigms utilizing One-Way ANOVA:

1. Is there a difference in Learning based on the type of instruction in the CCIS? (Comparative Design)

"Learning" Measure

A: Lecture

B: All ECR

C: Videotape

} Collect Data from each sample

2. Is time (i.e., acquaintance with the medium) a factor in the ability to learn through ECR instruction?
(Experimental Design)

"Measured" Ability to Learn

X_1 : Control

X_2 : Six months of ECR instruction

X_3 : One year of ECR instruction

Group	ECR Instruction	Post test
X_1	No	Yes
X_2	Yes	Yes
X_3	Yes	Yes

ANOVA can be determined by using either

1. Deviation Data
2. Raw Data

Deviation Data

1. Determine the deviation of each score from the mean

$$X - \bar{X}$$

2. Determine "within sample estimate of Variance"

$$MS_w = \frac{\sum (X - \bar{X})^2}{kn} = \frac{\sum (X - \bar{X})^2}{k(n-1)}$$

MS = Mean Square

MS = Mean square for the "within" sample estimate of variance

k = number of subsamples (independent) samples

n = number of scores in each subsample

Note. This formula is based on the assumption that the sample number is equal.

3. Determine the "estimate of variance" between or among the samples.

$$MS_b = \frac{(n) (\sum (\bar{X} - \bar{\bar{X}})^2)}{k-1}$$

$\bar{X}$ = mean of an individual sample

$\bar{\bar{X}}$ = grand mean

k = number of subsamples

n = number of scores per subsample

df = k-1 (for the estimate of between sample variance)

4. Determine F ratio

$$F_{ratio} = \frac{MS_b}{MS_w} \quad 80$$

Determine the ANOVA F statistic with Deviation Data:

X_1	X_2	X_3
2	4	6
4	6	8
6	8	10
Σ 12	18	24
$\bar{X}$ 4	6	8

$$\text{Grand Mean} = \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{K}$$

$$= (4 + 6 + 8) / 3 = 18 / 3$$

$$= 6, \text{ or } \bar{\bar{X}}$$

$$\text{Computed Mean} = \frac{\Sigma X_t}{N} = \frac{(12 + 18 + 24)}{9} = 6$$

$$MS_w = \frac{\sum_K \sum_n (\bar{X} - \bar{\bar{X}})^2}{K(n-1)}$$

$$= \frac{[(2-4)^2 + (4-4)^2 + (6-4)^2] + [(4-6)^2 + (6-6)^2 + (8-6)^2] + [(6-8)^2 + (8-8)^2 + (10-8)^2]}{3 \times (3-1)}$$

$$= \frac{[4 + 0 + 4] + [4 + 0 + 4] + [4 + 0 + 4]}{(3)(2)} = \frac{8 + 8 + 8}{6} = 4$$

$$df = k(n-1) = 3(3-1) = (3)(2) = 6$$

$$MS_b = \frac{(n) \left[\sum_k (\bar{x} - \bar{\bar{x}})^2 \right]}{k-1}$$

$$= \frac{(3) [(4-6)^2 + (6-6)^2 + (8-6)^2]}{3-1}$$

$$= \frac{(3)(4+0+4)}{2} = \frac{(3)(8)}{2}$$

$$= 24/2 = 12$$

$$df = k-1 = 3-1 = 2$$

$$F = \frac{MS_b}{MS_w} = \frac{12}{4} = 3.0$$

Variation	SS	df	MS	F
Between	24	2	12	3.00
Within	24	6	4	

ANOVA.02-05

$$F(\alpha = .05, 2, 6) = 5.14$$

$$F(\alpha = .01, 2, 6) = 10.92$$

$$\text{Computed } F = 3.00$$

$$\text{Computed } F (3.00) < 5.14 (\alpha = .05)$$

$$10.92 (\alpha = .01)$$

$\therefore$ No significant difference

Conclusion: There are no true differences among the three groups (X_1, X_2, X_3).

Differences are due to chance.

Determine the ANOVA F Statistic
with raw data :

X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
10	100	13	169	21	441
9	81	9	81	17	289
14	196	17	289	15	225
15	225	21	441	14	196
8	64	17	289	25	625
12	144	15	225	25	625
14	196	9	81	27	729
17	289	16	256	18	324
20	400	19	361	21	441
13	169	23	529	22	484
$n = 10$	$\Sigma = 1,864$		2,721		4,379
$\Sigma = 132$		159		205	

$$\Sigma X_t^2 = 1,864 + 2,721 + 4,379$$

$$\Sigma \bar{X}_t^2 = 8,964$$

$$\Sigma X_t = 132 + 159 + 205$$

$$\Sigma X_t = 496$$

ANOVA.02-07

$$1. SS_T = \sum X_t^2 - \frac{(\sum X_t)^2}{N} = 8,964 - \frac{(496)^2}{30}$$

$$= 8,964 - (246,016/30) = 8,964 - 8,200,533$$

$$= 763.467$$

$$df = N \cdot 1 = 30 \cdot 1 = 29$$

$$2. SS_W = \sum X_t^2 - \sum \frac{(\sum X_k)^2}{n_k}$$

$$= 8,964 - \left[\frac{(132)^2}{10} + \frac{(159)^2}{10} + \frac{(205)^2}{10} \right]$$

$$= 8,964 - \left[\frac{17,424}{10} + \frac{25,281}{10} + \frac{42,025}{10} \right]$$

$$= 8,964 - (1,742.4 + 2,528.1 + 4,202.5)$$

$$= 8,964 - 8,473$$

$$= 491$$

$$df = k(n-1)$$

$$= 3(10-1)$$

$$= 3(9)$$

$$df = 27$$

$$\begin{aligned}
 3. \quad SS_b &= SST - SSW \\
 &= 763.467 - 491 \\
 &= 272.47
 \end{aligned}$$

$$df = k - 1 = 3 - 1 = 2$$

$$4. \quad MS_b = \frac{SS_b}{k-1} = \frac{272.47}{2} = 136.24$$

$$5. \quad MS_w = \frac{SS_w}{k(n-1)} = \frac{491}{3(10-1)} = \frac{491}{3(9)} = 18.19$$

6. Develop the Standard ANOVA Table:

Source of Variation	df	SS	MS	F
Treatment (between)	2	272.47	136.23	7.49 **
Error (within)	27	491.00	18.18	
TOTAL	29	763.47		

$$7. \quad F = \frac{MS_b}{MS_w} = \frac{136.23}{18.18} = 7.49 **$$

Conclusion:

Computed $F(7.49) >$ Criterion $F(3.35; 5.49)$

$$\begin{aligned}
 (\alpha, 05, 2/27 &= 3.35) \\
 (\alpha, 01, 2/27 &= 5.49)
 \end{aligned}$$

∴ Differences are significant and not
Due to chance

Tips on Using ANOVA and the F test

- ~ it is assumed that random selection is in effect
- ~ yet the F test is often a valuable statistic even when random selection is not the case
- ~ but placement into subgroups shall be effective in the F test as to be a valid statistic
- ~ no difference will be the result if there is no difference between/among the various samples

Note: From agriculture to education, the F test and ANOVA are highly valuable tests to explain what is otherwise difficult to explain. ANOVA is often the statistical test of choice if:

1. Two or more samples are being reviewed against one or more variables.
2. Randomization was not fully effected. (But, every effort should be made to effect random selection).

ANALYSIS OF VARIANCE: Two Way

* the statistical test for two or more variables that each has two or more variables

~ USED to determine if:

1. there is a difference because of variables acting independently of each other, or
2. there is a difference because of joint effects

~ like chi-square Two Way ANOVA uses a factorial organization with data placed in cells

~ the information within each cell provides the necessary data for later analysis

An outcome of Two Way ANOVA is that:

1. The effect(s) due to application of the manipulated variables can be found

2. The interaction between variables can also be found

The F statistic is obtained for each source of variation.

Two Way ANOVA is common for comparative as well as experimental studies

Comparative Study of Two Variables,
with two categories for each variable.

		VARIABLE A	
		Category 1	Category 2
Variable 3	Category 1		
	Category 2		

Experimental Study of Two Variables,
with two categories for each variable.

H₀: There is no difference between
teaching methods (CBT, Lecture)
on learning outcomes among students

Independent Variables:

1. CBT
2. Lecture

CATEGORIES for each Variable:

1. One Hour of instruction between breaks
2. One-Half hour of instruction between breaks

Variable A: CBT

(1) 1 hour of instruction, break n=5

(2) .5 " " " " n=5

Variable B: Lecture

(1) 1 hour of instruction, break n=5

(2) " " " " n=5

A1	A2	B1	B2
17	42	15	29
27	29	17	36
15	31	19	29
19	35	12	47
12	36	23	34

Note: For this sample

1. It is assumed that an acceptable Random selection process was used and that subjects were not aware of their participation in a study.
2. Outcomes data are from posttest results. The Associated posttest is regarded as reliable / valid.

ANOVA.03-03

	1.00 Hour	0.50 Hour	
CBT	A1	A2	
	17	42	
	27	29	
	15	31	
	19	35	
	12	36	
	$\Sigma = 90$	$\Sigma = 173$	$\Sigma Row = 263$
	$\bar{X} = 18.0$	$\bar{X} = 34.60$	
Lecture	B1	B2	
	15	29	
	17	36	
	19	29	
	12	47	
	23	34	
	$\Sigma = 86$	$\Sigma = 175$	$\Sigma Row = 261$
	$\bar{X} = 17.2$	$\bar{X} = 35.0$	
	$\Sigma C = 176$	$\Sigma C = 348$	

Process for Calculating the Appropriate F Statistics:

1. Add the scores for each cell

$$A1 = 90$$

$$A2 = 173$$

$$B1 = 86$$

$$B2 = 175$$

$$\begin{aligned}
 2. \quad \Sigma \text{ Overall} &= A_1 + A_2 + B_1 + B_2 \\
 &= 90 + 173 + 86 + 175 \\
 &= 524
 \end{aligned}$$

$$\begin{aligned}
 3. \quad \text{Determine column totals} \\
 \Sigma C &= 176 + 348 \\
 &= 524
 \end{aligned}$$

$$\begin{aligned}
 4. \quad \text{Determine row totals} \\
 \Sigma R &= 263 + 261
 \end{aligned}$$

5. Determine sum of squares for each scene in the overall set of Data.

A_1	A_1^2
17	289
27	729
15	225
19	361
12	144
Σ	1,748

A_2	A_2^2
42	1,764
29	841
31	961
35	1,225
30	1,296
Σ	6,087

B_1	B_1^2
15	225
17	289
19	361
12	144
23	529
Σ	1,548

B_2	B_2^2
29	841
30	1,296
29	841
47	2,209
34	1,156
Σ	6,343

$$\Sigma X_t^2 = 1,748 + 6,087 + 1,548 + 6,343$$

$$\Sigma X_t^2 = 15,726$$

PROVA.03-05

5. Clearly identify all symbols for
later use:

$$\sum A_1 = 90$$

$$\sum A_2 = 173$$

$$\sum B_1 = 86$$

$$\sum B_2 = 175$$

$$\sum X_t^2 = 15,726$$

$$N = 20 \quad \dots \text{total number of samples}$$

$$n_k = 5 \quad \dots \text{number of scores per cell}$$

$$R = 2 \quad \dots \text{number of rows}$$

$$C = 2 \quad \dots \text{number of columns}$$

$$k = 4 \quad \dots \text{number of cells}$$

$$n_c = 10 \quad \dots \text{number of samples in columns}$$

$$n_r = 10 \quad \dots \text{number of samples in rows}$$

6. Determine the total sum of squares

$$SS_t = \sum X_t^2 - \frac{(\sum X_t)^2}{N}$$

$$SS_t = 15,726 - \frac{(90+173+86+175)^2}{20}$$

$$SS_t = 15,726 - \frac{(524)^2}{20}$$

$$SS_t = 15,726 - \frac{274,576}{20}$$

$$SS_t = 15,726 - 13,728.80$$

$$SS_t = 1,997.20$$

$$df(SS_t) = N - 1$$

$$df(SS_t) = 20 - 1$$

$$df(SS_t) = 19$$

7. Determine "within cell variance"
sum of squares

$$SS_w = \sum X_t^2 - \sum \frac{(\sum X_k)^2}{n_k}$$

$$SS_w = 15,726 - [(90)^2/5 + (173)^2/5 + (86)^2/5 + (175)^2/5]$$

$$SS_w = 15,726 - (1620.0 + 5985.8 + 1479.2 + 6125.0)$$

$$SS_w = 15,726 - (15,210)$$

$$SS_w = 516$$

$$df(SS_w) = k(n-1)$$

$$df(SS_w) = 4(5-1)$$

$$df(SS_w) = 16$$

8. Determine ^{between} ~~within~~ column sum of squares

$$SS_c = \sum \frac{(\sum X_c)^2}{n_c} - \frac{(\sum X_t)^2}{N}$$

$$SS_c = [(176)^2/10 + (348)^2/10] - \frac{(524)^2}{20}$$

$$SS_c = (3077.6 + 12,110.4) - 13,728.8$$

$$SS_c = 15,208 - 13,728.8$$

$$SS_c = 1,479.2$$

$$df(SS_c) = C - 1$$

$$df(SS_c) = 2 - 1$$

$$df(SS_c) = 1$$

9. Determine "between row"
sum of squares

$$SS_r = \sum \frac{(\sum X_r)^2}{n_r} - \frac{(\sum X_t)^2}{N}$$

$$SS_R = [(263)^2/10 + (261)^2/10] - (524)^2/10$$

$$SS_R = [6,916.9 + 6,812.1] - 13,728.8$$

$$SS_R = 13,729.0 - 13,728.8$$

$$SS_R = 0.20$$

$$df(SS_R) = 12 - 1$$

$$df(SS_R) = 2 - 1$$

$$df(SS_R) = 1$$

10. Determine "between cell" sum of squares:

$$SS_b = SS_t - SS_w$$

$$SS_b = 1,997.2 - 516.0$$

$$SS_b = 1,481.2$$

$$df(SS_b) = k - 1$$

$$df(SS_b) = 4 - 1$$

$$df(SS_b) = 3$$

ANOVA.03-09

11. Determine 'between variables' sum of squares
~ interaction

$$SS_I = SS_b - (SS_c + SS_r)$$

$$SS_I = 1,481.2 - (1,479.2 + 0.2)$$

$$SS_I = 1,481.2 - (1,479.4)$$

$$SS_I = 1.8$$

$$df(SS_I) = (D-1)(C-1)$$

$$df(SS_I) = (2-1)(2-1)$$

$$df(SS_I) = 1$$

12. Determine the mean square for each sum of squares, except SS_I

$$MS = \frac{SS}{df}$$

$$MS_w = \frac{SS_w}{df_w} = \frac{516}{16} = 32.25$$

$$MS_c = \frac{SS_c}{df_c} = \frac{1,479.2}{1} = 1,479.2$$

$$MS_R = \frac{SS_R}{df_R} = \frac{0.2}{1} = 0.2$$

$$MS_I = \frac{SS_I}{df_I} = \frac{1.8}{1} = 1.8$$

$$MS_b = \frac{SS_b}{df_b} = \frac{1,481.2}{3} = 493.73$$

13. Develop an ANOVA table to determine F ratios

Source of Variance	SS	df	ms	F
Between Cells	1,481.2	3	493.7	15.31**
between rows	0.2	1	0.2	0.01
" columns	1,479.2	1	1,479.2	45.87**
Interaction	1.8	1	1.8	0.06
Within Cells (Error)	516.0	16	32.25	
Total	1,997.2	19		

Note: Between Cells SS = $SS_R + SS_c + SS_I$

$$F_{\text{cells}} = \frac{MS_{\text{B}}}{MS_{\text{W}}} = \frac{493.7}{32.25} = 15.308$$

$$F_{\text{Rows}} = \frac{MS_{\text{R}}}{MS_{\text{W}}} = \frac{.2}{32.25} = 0.01$$

$$F_{\text{columns}} = \frac{MS_{\text{C}}}{MS_{\text{W}}} = \frac{1,479.2}{32.25} = 45.867$$

$$F_{\text{Interaction}} = \frac{MS_{\text{I}}}{MS_{\text{W}}} = \frac{1.8}{32.25} = 0.06$$

$$\text{Criterion } 1/16 (d_{.05}) = 4.44$$

$$\text{Criterion } 1/16 (d_{.01}) = 8.53$$

Conclusion:

There is a significant difference in the effect of time between breaks (between columns). [$F_{\text{C}} = 45.867 **$]

But, there is no significant difference between teaching methodology or interaction.

Conclusions:

1. ~~Major~~ difference relates to time between breaks (1.0 hr as opposed to .50 hours)
2. The use of CST or Lecture (and subsequent interaction vs. time) has no effect on learning outcomes

Thus - Time between breaks is the only phenomenon of importance.